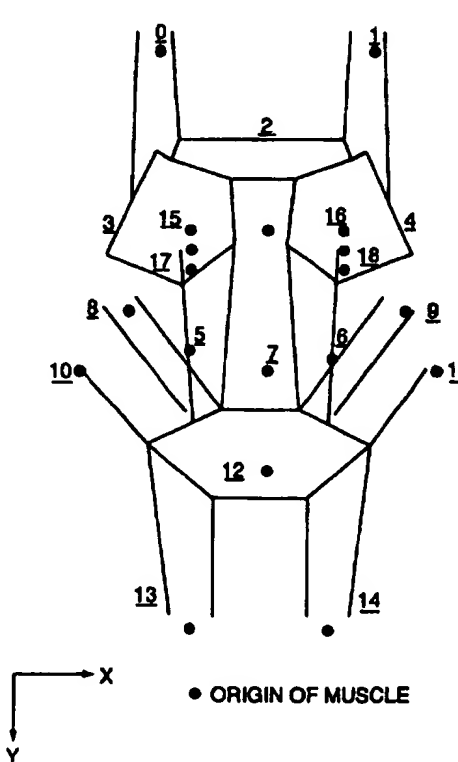




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : <b>G06T 15/70</b></p>	<b>A1</b>	<p>(11) International Publication Number: <b>WO 98/01830</b></p> <p>(43) International Publication Date: <b>15 January 1998 (15.01.98)</b></p>
<p>(21) International Application Number: <b>PCT/GB97/01834</b></p> <p>(22) International Filing Date: <b>7 July 1997 (07.07.97)</b></p> <p>(30) Priority Data:              9614194.0           5 July 1996 (05.07.96)           GB              96306184.1       23 August 1996 (23.08.96)       EP          (34) Countries for which the regional or international application was filed: <b>GB et al.</b></p> <p>(71) Applicant (for all designated States except US): <b>BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).</b></p> <p>(72) Inventors; and          (75) Inventors/Applicants (for US only): <b>WELSH, William, John [GB/GB]; 249 South Lane, New Malden, Surrey AT3 5RP (GB). PINCEMAIL, Cecile [FR/FR]; 65 rue Ernest Renan, F-92310 Sevres (FR).</b></p> <p>(74) Agent: <b>HIBBERT, Juliet, Jane, Grace; BT Group Legal Services, Intellectual Property Dept., 8th floor, 120 Holborn, London EC1N 2TE (GB).</b></p>		
<p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</b></p> <p><b>Published</b>  <i>With international search report.</i></p>		
<p>(54) Title: <b>IMAGE PROCESSING</b></p> <p>(57) Abstract</p> <p>A method of coding an image of an animated object, an object being represented by a shape model defining the generic shape of the object and a muscle model defining the generic arrangement of muscles associated with the object, said image being coded in terms of movement of the shape and/or muscle model, the shape and muscle model for an object having a predefined interrelationship such that when one of the models is conformed to the shape of a specific example of the object, the other of said models is also conformed accordingly. The muscle model comprises information relating to predefined expressions, which information relates to which muscles are activated for each predefined expression and the degree of activation required, wherein, when the shape model is conformed to an object, the degree of activation is adapted in accordance with the changes made to the shape model.</p>		
		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## IMAGE PROCESSING

This invention relates to image processing and in particular image processing for use in real time applications such as video conferencing.

5        Broadcast quality TV requires a transmission bandwidth in excess of 100 Mbit/s when coded in digital form which is both expensive to transmit and requires high bandwidth links. When images are coded digitally for transmission, it is therefore desirable to keep the amount of data generated as low as possible.

      This may be done using predictive coding techniques which exploit the  
10 correlation between the picture elements (pixels) of the digitised sequence of images or frames. Alternatively model-based or knowledge-based image coding may be used in which a model of an object in a scene is used. As an example, consider a scene containing a cube moving around against a plain background: the information needed to generate a moving sequence at the receiver would be a  
15 description of the cube (which only needs to be sent once at the beginning of the transmission) followed by a description of its motion.

      It is very straightforward to produce computer animated sequences of cubes using model-based coding, but for video conferencing the modelling of people in the scene is more complicated. In "Parameterised Models for Facial  
20 Animation" FI Parke, IEEE Computer Graphics and Applications, November 1982, a parameterised model of a face was proposed for use in videophony. A computer model of a head comprising a net of inter-connected polygons is stored in a computer as a set of linked lists or links arrays.

      In order to make the models appear more realistic, the polygon net can be  
25 shaded. The shading is made to depend on the presence of imaginary light sources in the model world and the assumption of reflectance properties of the model surfaces. A so-called smooth shading technique such as Gouraud or Phong shading can be used to improve realism. In these techniques the shading of each polygon is effectively interpolated from the shading values at the vertices. This  
30 gives the impression of a smoothly curved surface.

      The face can be made to move in a global way by applying a set of rotations and translations to all the vertices as a whole. In addition the vertices of certain polygons can be translated locally in order to change the facial expression.

This may be in accordance with a system that has been developed for categorising facial expressions in terms of about 50 independent facial actions (the facial action coding system or FACS) as described in "Manual for the Facial Action Coding System" P Ekman, W V Friesen, Consulting Psychologists Press, Palo Alto, California 1977.

The diversity of facial forms in terms of sex, age and race is enormous. It is these diversities that allow us to recognise individuals and send complex non-verbal signals to one another. However, the anatomy of a human head is the same for everyone. It differs in size, but the structure, the arrangement of the bones and their rules are the same. Generally everyone has got the same bones in the same place and the same muscles.

In accordance with the invention there is provided a method and apparatus for coding an image of an animated object, an object being represented by a shape model defining the generic shape of the object and a muscle model defining the generic arrangement of muscles associated with the object, said image being coded in terms of movement of the shape and/or muscle model, the shape and muscle model for an object having a predefined interrelationship such that when one of the models is conformed to the shape of a specific example of the object, the other of said models is also conformed accordingly.

Thus this generic structure of a shape and a muscle model can be easily manipulated for a specific object.

Preferably the muscle model comprises information relating to predefined expressions, which information relates to which muscles are activated for each predefined expression and the degree of activation required, wherein, when the shape model is conformed to an object, the degree of activation is adapted in accordance with the changes made to the shape model.

The generic shape model and the generic muscle model may share common points within an object.

A set of action units may be generated for a specific object, each action unit defining the displacement of points within the specific object necessary to produce a required animated expression.

The invention is particularly suitable for use in video conferencing applications and will also find use in graphical user interfaces which require animation.

The invention will now be described further by way of example only with  
5 reference to the accompanying drawing in which:

Figure 1 shows an example of a generic shape model for a human face;

Figure 2 shows an example of a generic muscle model for a human face;

Figure 3 shows the logical structure of the muscle model of Figure 2;

Figure 4 shows a structure of the expressions associated with the muscle  
10 model;

Figure 5a and b show the generic shape model of Figure 1 and the conformed shape model respectively;

Figure 6 shows a conformed muscle model added onto a face image; and

Figures 7 a-f show various expressions applied to the face.

15 The shape model, as shown in Figure 1, represents a human head and comprises a so-called wire frame of interconnected polygons 10. The x, y and z components of each polygon vertex in the object's face are stored in a computer as arrays  $X(V)$ ,  $Y(V)$ ,  $Z(V)$  where  $V$  is the vertex address. In addition, a pair of two-dimensional arrays are used:  $LIN\ V(L)(E)$  gives the addresses of the vertices at  
20 the end of a line  $L$ , where  $E$  is either 0 or 1 depending on which end of the line is being considered:  $LIN\ L(P)(S)$  gives the line of address for each side  $S$  of a polygon  $P$ . The wire frame is drawn on the screen by iterating through all the values of  $L$  in  $LIN\ V(L)(E)$  giving the vertex addresses which, in turn, yield the co-ordinates of the vertices using the arrays  $X(V)$ ,  $Y(V)$  and  $Z(V)$ . The side of the polygon is projected  
25 onto the screen using perspective or orthographic projection as is known in the art of model-based image coding.

The muscles of the human face are embedded in the superficial fascia, the majority extending from the bones of the skull and into the skin. The orifices of the face, namely the eye orbit, nose, and mouth, are guarded by the eyelids, nostrils  
30 and lips. It is the function of the facial muscles to serve as sphincters or dilators of these structures. A secondary function of the facial muscles is to modify the expression of the face.

Each muscle can be defined according to the orientation of the fasciculi (the individual fibres of the muscles) that may be parallel/linear, oblique or spiralized relative to the direction of pull at their attachment. They can be split into two groups: the upper and lower face.

- 5        In the lower face there are five major groups:
- Uppers and downers, that move the face upwards towards the brow and conversely towards the chin
  - Those that contract horizontally towards the ears and conversely towards the centre line of the face
  - 10      • Oblique muscles that contract in an angular direction from the lips, upwards and outwards to the cheek bones
  - Orbitals, that are circular or elliptical in nature, and run round the eyes and mouth
  - Sheet muscle, that carries out miscellaneous actions, particularly over
  - 15      the temporal zones

The upper facial muscles are responsible for the changing appearance of the eyebrows, forehead and the upper and lower lids of the eyes. The muscles contract isotonicly towards the static insertion into the cranium, consequently the surface tissue bunches and wrinkles perpendicularly to the direction of the

20    muscle.

The muscles of the mouth have the most complex muscular interaction. The Obicularis Oris is a sphincter muscle with no bony attachment.

The facial muscle system is very complex. For videoconferencing purposes a very complicated model was undesirable due to the small size of the image and the real time nature of a videoconferencing application. A simple muscle model for

25    the face is here described which, for the sake of simplicity only, does not include any muscles on the mouth.

Figure 2 shows an example of a generic muscle model for the muscles of the face. In Figure 2 the origin of each muscle is represented by a solid circle.

- 30      • **Muscles 0 and 1** represent the frontalis muscle, which are used, for example, during the raised eyebrow expression. Each muscle is composed of two lines to represent the width of the frontalis muscle. As these muscles are attached to a bone, the upper vertices are not

allowed to move, and the origin is located between the two upper vertices of each muscle.

- 5       • **Muscle 2** represents the corrugator muscle. When it is contracted, the surface tissue bunches and wrinkles perpendicularly to the direction of the muscles. This muscle is composed by two segments whose extremities are moving vertices. The origin is positioned under the muscle to emphasize the slope of the eyebrow for an expression like anger.
- 10       • **Muscles 3 and 4** represent the orbicularis oculi, which are the sphincter muscles of the eyelids. In the model, each muscle describes a large circle around the eye to improve the realism of each motion. Each muscle contains 7 vertices, but just the upper vertices are allowed to move to represent the movement of the eyelid. The origin is in the middle of the circle.
- 15       • **Muscles 5 and 6** comprise one segment which creates a link between the eye and the mouth. This is used, for example, during a smile to contract the skin between the eye and the mouth.
- 20       • **Muscle 7** is the muscles of the nostrils used as a sphincter muscle and as a dilator muscle too. It contains two segments alongside the nose. The two lower vertices are used to represent the different motions of the nostrils. The origin is the middle point of the four extremities of the segments.
- 25       • **Muscles 8 and 9** can be compared to the zygomaticus major, and are very useful for the smile. Each muscle is represented by two segments to show the width of the muscle. To translate the attachment to the bone, only the lower vertices are allowed to move, and the origins are between the two upper extremities.
- 30       • **Muscles 10 and 11** can be compared, at one and the same time, to the buccinator and the m levator. Attached to the bone on one extremity, just the nearest to the mouth vertices are allowed to move.

- **Muscle 12** is the obicularis oris. This is a sphincter muscle with no bony attachment so each vertex is allowed to move. This muscle contains 8 segments and is used in most of the expression. It is not used only as a sphincter muscle which compresses the lips together, but as a dilator muscle which also separates the lips. It just defines the shape of the external mouth. The internal mouth is not modified by this muscle.
- **Muscles 13 and 14** can be compared, at one and the same time, to the M depressor and to the labii inferioris. They contain two segments to represent the width of the muscle, and two moving vertices to represent the bony attachment.
- **Muscles 15 and 16** are used to move the eyelids. By their contraction the accuracy is better than by the orbicularis oculi. One the single segment, the upper vertex is used as origin and the lower is allowed to move.
- **Muscles 17 and 18** are used to represent the small contraction of the skin under the eyes. As the two previous muscles, the accuracy is better. They are composed on the same base.

The face wire frame as shown in Figure 1 is modified to add new vertices defining the muscles' extremities i.e. the muscle model of Figure 2 is added to the shape model of Figure 1. The face wire frame is a set of 144 vertices, and some of them are used as muscle points too. 54 other points are defined as muscles' extremities only because no vertex in the shape model corresponds to the position of the muscle extremity. The information is stored in a file. The first line of this file contains the identification number of a face wire frame (e.g. no. 0001), then the number of vertices in the face wire frame (144), and before the number of facets (217), there is the number of vertices used only as muscles' extremities (54).

The muscle model and the shape model are locked together such that, when the generic shape model is conformed to fit a particular face, the muscle model is also conformed. As well as positioning the muscle attachment points this also scales the muscle lengths. When an expression is used to activate a set of muscles - for example, to create an impression of anger - the muscles thus



contract to an appropriate degree for the specific face, rather than for the generic face where different sizes and locations of the features (eyes, mouth, nose etc.) would require a different set of characteristics.

The logical structure of the muscle model is defined as a tree, as shown in  
5 figure 3.

On the top level 31, there is the muscle model of the face (as shown in Figure 2) which represents the set of all the muscles of the face, which are linked together. This set is characterised by the number of muscles 31a that it contains and by a list 31b of all the individual muscles.

10 Each muscle is composed of one or more segments (1 to n). For example the Orbicularis Oris, which is around the mouth, contains 8 segments. Each segment is defined by the vertices defining its extremities. These vertices belong to the face wire frame as will be described later. The segments of each muscle are defined in the next level 32.

15 Two other parameters used to characterise a muscle are:

- the origin 32a, which is used when the muscle is pulled or squeezed. The origin is defined by x, y, z co-ordinates or a vertex number which define(s) the centre of the scale function corresponding to the muscle activation (the origins of the muscle model are indicated by dots in  
20 Figure 2).
- the moving vertices 32b, which characterised the extremities of the muscle attached to the skin (i.e. the position of the moving vertices). Those vertices which represent muscle extremities which are attached to the bones of the skull and hence which do not move are not defined  
25 in the muscle model.

All the muscles are described in a computer file by their segments 32, moving vertices 32b and origin 32a. This file contains one line per muscle. For each muscle, there is the number of the segment, the list of the segment extremities; then the number of moving vertices and the list of moving vertices,  
30 and the number and the list of vertices used to calculate the position of the origin. The origin is the barycentre of these vertices.

Each expression is described by a list of muscles to contract. In a file named *expression*, each expression is defined in terms of the number of muscles

to scale, the muscle number identification, and the X and Y factors by which the muscles are to be expanded or contracted with respect to the origin to produce the desired expression.

The logical structure of the expressions is shown in figure 4.

5       The X and Y factors determine the displacement (in the X and Y directions) of the moving vertices of the associated muscle with respect to the origin. For instance, consider muscles 5 and 6 of Figure 2, from the eye to the mouth. The generic muscle model may require these muscles to contract by 30% in the Y direction and 2% in the X direction for the expression of a smile. Thus  
10   the X and Y factors for the expression "smile" for the muscles 5 and 6 are -2 and -30 respectively (or 98% and 70% of the initial position with respect to the origin).

Rather than defining the muscle contraction factors as factors of X and Y, the muscle contraction factors may represent the proportion by which the muscle length has to contract for a given expression, the orientation of the moving  
15   vertices with respect to the origin remaining the same.

The above has assumed that the degree of activation of the muscle is defined once for each muscle used in an expression. Alternatively, however, the degree of activation may be defined for each vertex of the muscle.

Thus the generic muscle model comprises a set of muscle definitions (as  
20   shown logically in Figure 3) and a set of expressions (as shown logically in Figure 4). To convert these expressions to so-called action units, the initial position of each point of the face wire frame is registered, and the position when the expression is maximal. Then a program determines the difference between the positions and stores them as a action unit. The last step consists of adding those  
25   lines to the action unit file.

As for the shape wire frame, this generic muscle model has to be manipulated for a specific head. The picture in Figure 6 shows a manipulated muscle model added onto a face image, as will be described below.

To conform the shape and muscle models to a particular user, firstly at  
30   least two images of the person are grabbed by a video camera, including one image with the mouth open and the other with the mouth closed. Both of these images are loaded into a computer, and the generic shape and muscle models are adapted by the conform program, from the generic model, to the specified person.

The conform program is also used to define and test all the action units involved in the animation. Once the tests are realised, the files are send to a receiver. Then the person at the transmitter is able to speak on demand via the adapted talking head at the receiver.

5       The head is produced from two still pictures, one with the mouth closed, the other with the mouth open. These pictures have the format \*.yuv. Loading the data files means defining the wire frames. So this part is divided into two sections: work on the face wire frame with the mouth closed image and work on the mouth wire frame with the mouth open image.

10       First, to adapt the shape wire frame for the head, the *face.yuv* file is loaded together with the generic face wire frame. Figure 5a shows an image of a participant with the generic shape model superposed thereon. The wire frame is then scaled and rotated by an operator to conform the generic model to the object. An example of the conformed shape model is shown in Figure 5b.

15       Then the conformed wire frame can be made more accurate by the use of key points. The key points are special vertices of the wire frame which can be moved by the mouse. A triangulation is made from all the key points; moving one key point modifies one or more triangles and moves all the vertices of the wire frame contained in those triangles. So moving the key points allow a better  
20 accuracy for the wire frame.

The same actions are then carried out for the mouth wire frame (not shown).

The details sent to the receiver comprise six files:

- 25
- Two files containing the two pictures used, named *face.fy*, for the picture with the mouth closed, and *mouth.my*, for the picture with the mouth open. These pictures have the same size as the talking head window.
  - Two files describing the vertices of each wire frame, *face.fv* for the face wire frame and *mouth.mv* for the mouth wire frame.
  - 30   • Two files for the facets of the wire frame, *face.ff* and *mouth.mf*.

All these files are created by the conformation program from the pictures produced by the video, and from the manipulated wire frames. Then they are used

to produce an animated image at the transmitter, which is then used by the talking head at the receiver.

After conformation of both the generic wireframe model and the generic muscle model to a specific face, synthesised facial expressions can be produced on the face. This can be achieved by applying the set of contraction factors to certain muscles as explained with reference to Figure 4. This has two effects: first, other attached muscles are tugged along since all the muscle form a linked set; second, the positions of the wireframe vertices are affected by the muscle movements. The method used to move the wireframe vertices is the same as that used during the conformation process, in which some of the vertices are designated as keypoints which the user can reposition and the remaining vertices are moved automatically using the interpolation process as described in relation to the shape model. In the case of movement due to muscle action, the endpoints of the muscles take the place of the keypoints in the conformation process and the same interpolation method is used.

The method described to produce the synthesised facial expressions may be used on any face. To determine one set of muscle contraction factors which produce satisfactory facial expressions for most faces, it is necessary to consider a plurality of different faces to obtain average factors which produce a desired result. Thus a head independent set of muscle contraction factors is obtained. This method of determining muscle contraction factors is used off-line since it is more computationally intensive than the method of using action units because of the interpolation process. The method of action units is fast because it is based on a set of displacements applied to a subset of the vertices of the original wireframe model; only three additions or subtractions (in x, y, z directions) are required per vertex. Interpolation requires a number of multiplications to be carried out per vertex.

The same set of muscle contraction factors can then be applied to any new face for each of the expressions required and the displacement of each vertex of the wireframe model from the neutral position is recorded. The set of displacements obtained is then identical to an action unit which, if applied to a real-time player, will result in exactly the same expression being generated. Because the action unit is composed of a set of absolute displacements of the

wireframe vertices, it is specific to a particular head, unlike the set of muscle contractions which is generic. The action unit can be used in a real-time head player.

Although the preferred embodiment of the invention uses muscle  
5 contraction factors, the degree of activation of a muscle may alternatively be stored in absolute pixel values i.e. as so-called action units. In this case, the degree of activation may be adapted, in the X direction, in dependence on the ratio of the width of the generic shape model and the conformed shape model and similarly in the Y direction, in dependence on the ratio of the length of the two  
10 models. Alternatively, those vertices of the shape model which comprise areas of the object which are associated with muscles may be pre-defined and the relative displacement of the vertices (either individually or on average for a given muscle) of the generic and conformed shape model used to adapt the degree of activation.

An operator at a transmitter may check that the thus conformed animated  
15 head provides a good enough representation of the specific person. To do this each expression is tested. The operator is provided with a list of all the possible expressions. To activate an action unit, a cursor for the corresponding action is selected and the degree of the action unit to be applied adjusted if necessary.

Figures 7 a-f show different expressions produced using a generic  
20 muscle model as described:

- Figure 7a shows a smile. For a smile 10 muscles are activated. The most important contractions are the zygomatic contractions. However, the mouth is moving and stretching towards the side. The smile has repercussion for the eyes too.
- Figure 7b shows surprise. This expression activates 11 muscles. Most of the modifications are around the eyes to open them. Then the muscle along the nose and the obicularis oris around the mouth are stretched down.

- Figure 7c shows anger. This activates 10 muscles: the muscles near the eyes for frowning, the muscle along the nose to enlarge the nostrils and the muscle on the chin to contract the skin on it.
- 5     • Figure 7d shows disgust. It involves 9 muscles. The same muscles as anger expression for frowning, but with a less degree, are activated. However, the most important transformations concern the scale of the mouth.
- 10    • Figure 7e shows a raised eyebrow. This expression concerns just 5 muscles near the eyes. The transformations are exactly the same as for surprise expression but limited to the upper part of the face and with a higher degree of application.
- Figure 7f shows sadness. This expression involves 6 muscles. Four muscles are used to lower the corner of the mouth. The two others are used to raise up the eyebrows.

15       Once those actions are done, the vertices are moved (if necessary), the displacements coded and the texture mapping is activated to show the animation. Several action units can be activated together.

      In the Figures, the muscle model is implemented on the face wire frame only, and displayed on the picture with the mouth closed. In practice, the mouth  
20    would also be animated but this has not been discussed here for the sake of simplicity.

      At this point, the muscles' vertices are defined for a particular user, but they need to be stored for the particular user in the structure defined in Figure 3. This operation is done when the face wire frame is loaded, by running a procedure  
25    named *read\_muscle*. This procedure uses the face wire frame and the file describing the muscle, *muscle.wfm*. Two other procedures were written to display the mesh of the muscles and their origins, *disp\_muscle\_wf* and *disp\_muscle\_origin*.

      When an operator modifies the muscle model, the muscles' extremities may be loaded as key points and a triangulation run. Displacing one extremity,

with the left button of a mouse, will modify the vertices contained in the modified triangles. Alternatively a muscle extremity may be selected with the right button of the mouse and then moved. This modification will have no repercussion on the whole face wire frame. In the same way, an origin of a muscle may be moved to  
5 adjust the centre of the scaling function.

It is possible to select an expression with an amount ranging from 0 to 100. This will run all the muscles' activation on X and Y axes. An amount of 100 will activate the contraction of the muscles with the pre-defined factor; an amount of 50 will activate the contraction by half as much, etc. Once the scale function is  
10 activated for all the muscles concerned by the expression, the conformation is loaded and the texture mapping is activated.

The degree of activation of a muscle for a given expression is defined generically. That means that the same vertices are moved for a smile whoever the person. However the amount of movement is adapted for a particular person by  
15 means of the conforming of the muscle model in accordance with the shape model.

The action units and images may be transmitted from a transmitter to a remote receiver, e.g. in a video-conferencing application. Alternatively, the action units and images may be used locally to provide an animated interface for a service  
20 or the like.

Whilst this description has described the invention in relation to an object representing a head, it will be clear to the reader that the invention is applicable to the animation of any object.

**CLAIMS**

1. A method of coding an image of an animated object, an object being represented by a shape model defining the generic shape of the object and a  
5 muscle model defining the generic arrangement of muscles associated with the object, said image being coded in terms of movement of the shape and/or muscle model, the shape and muscle model for an object having a predefined interrelationship such that when one of the models is conformed to the shape of a specific example of the object, the other of said models is also conformed  
10 accordingly, the muscle model also comprising information relating to predefined expressions, which information relates to which muscles are activated for each predefined expression and the degree of activation required, wherein, when the shape model is conformed to an object, the degree of activation is adapted in accordance with the changes made to the shape model.  
15
2. A method according to claim 1, wherein the generic shape model and the generic muscle model share at least one common point within an animated object.
3. A method according to claim 1 or 2, wherein a set of action units are  
20 generated for a specific object, each action unit defining the displacement of points within the specific object necessary to produce a required animated expression.
4. A method according to claim 1 or 2 wherein the degree of activation is represented as a proportion of the length of the associated muscle.  
25
5. An image reproduced from a signal coded according to any of claims 1 to 4.
6. An image processing apparatus comprising a stored shape model defining  
30 the generic shape of an object and a stored muscle model defining the generic arrangement of muscles associated with the object, means for coding an image of the object in terms of movement of the shape and/or muscle model, the shape and muscle model for an object having a predefined interrelationship such that when



one of the models is conformed to the shape of a specific example of the object, the other of said models is also conformed accordingly, the muscle model comprising information relating to predefined expressions, which information relates to which muscles are activated for each predefined expression and the  
5 degree of activation required, wherein, when the shape model is conformed to an object, the degree of activation is adapted in accordance with the changes made to the shape model.

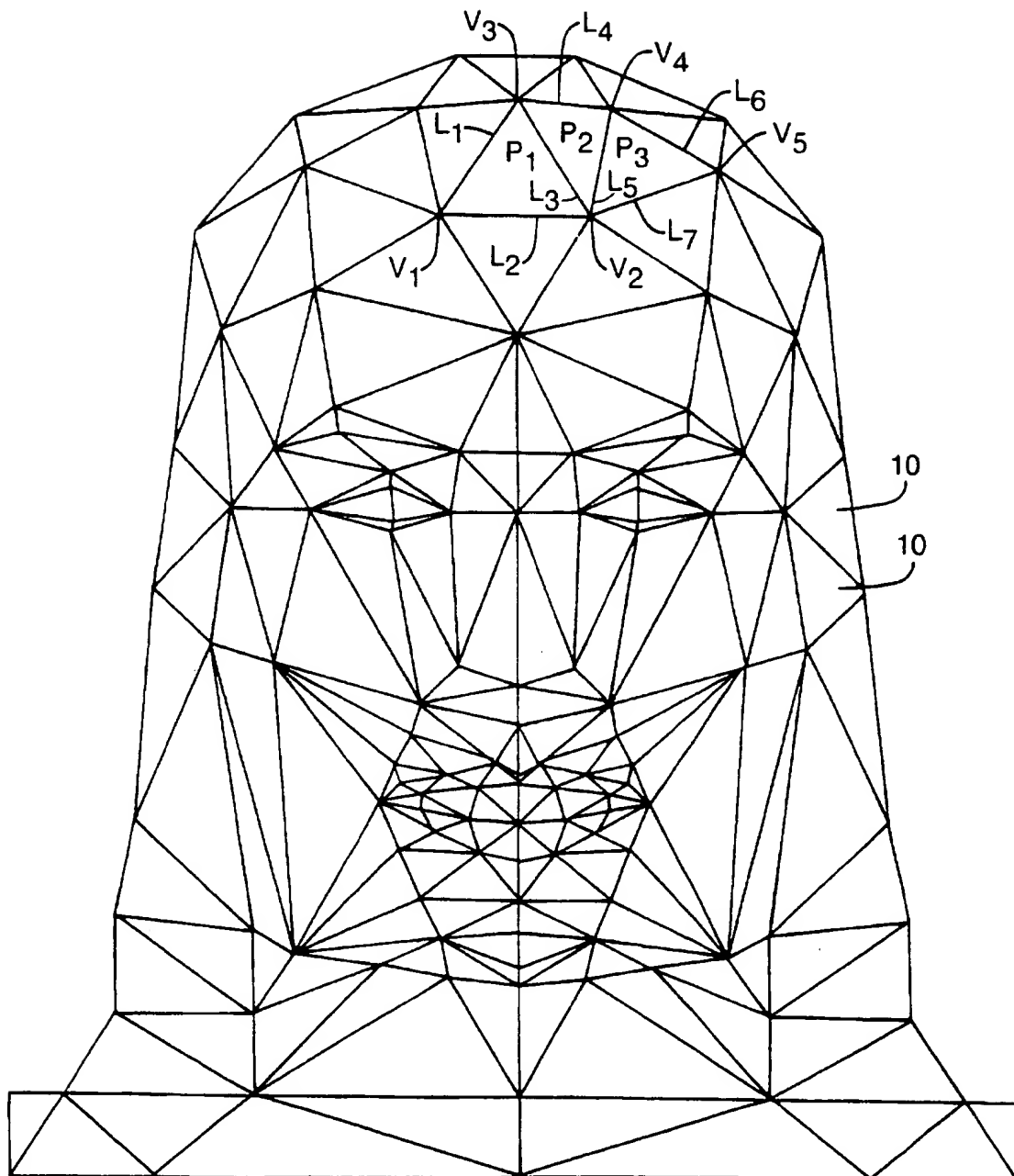
7. Apparatus according to claim 6, wherein the generic shape model and the  
10 generic muscle model share at least one common point within an animated object.

8. A method according to claim 6 or 7, wherein a set of action units are generated for a specific object, each action unit defining the displacement of points within the specific object necessary to produce a required animated expression.  
15

9. Video conferencing apparatus including image processing apparatus according to any of Claims 6 to 8.

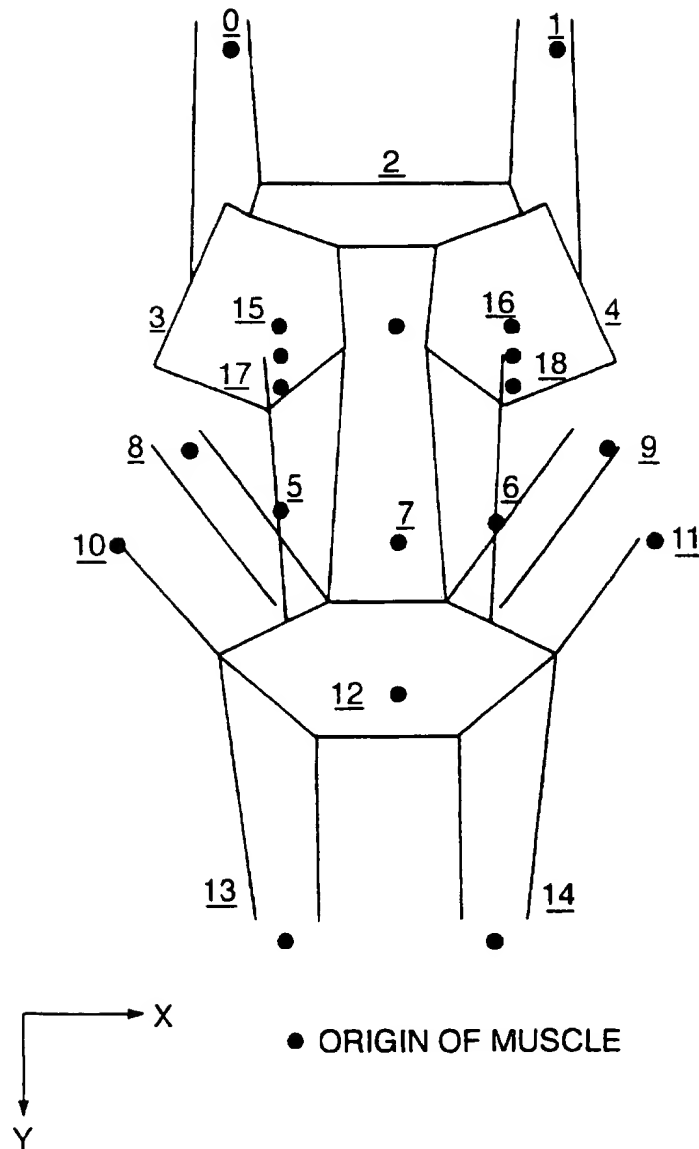
1/7

Figure 1



2/7

Figure 2



3/7

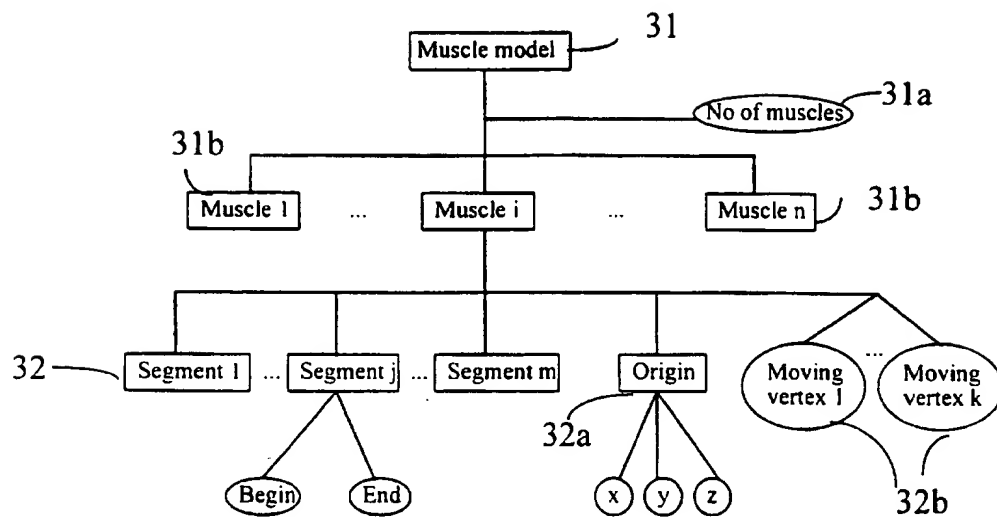


Figure 3

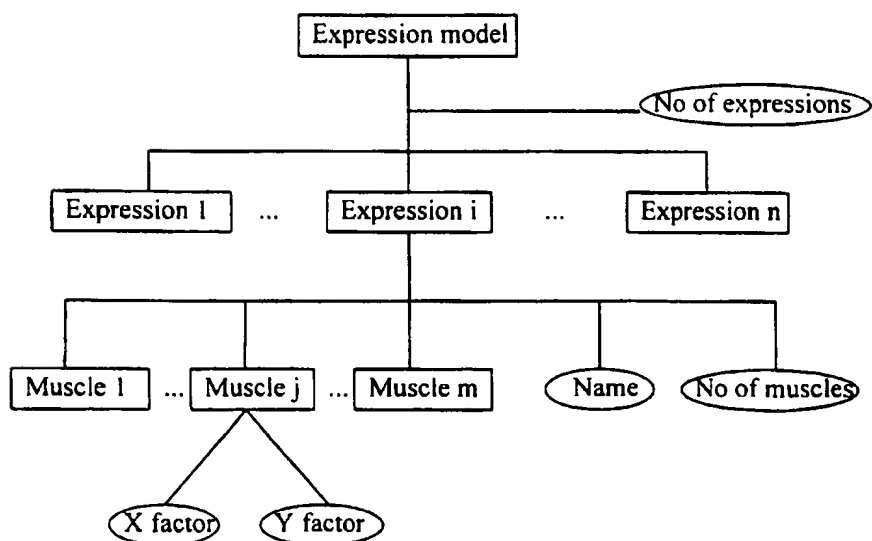


Figure 4

4/7



Figure 5a



Figure 5b



Figure 6

6/7



Figure 7a



Figure 7b



Figure 7c

7/7



Figure 7d



Figure 7e



Figure 7f



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 97/01834

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G06T15/70

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>YUENCHENG LEE ET AL: "REALISTIC MODELING FOR FACIAL ANIMATION" COMPUTER GRAPHICS PROCEEDINGS, LOS ANGELES, AUG. 6 - 11, 1995, 6 August 1995, COOK R, pages 55-62, XP000546216 see page 56, left-hand column, line 42 - line 50 see page 57, right-hand column, line 25 - line 32 see page 58, right-hand column, line 3 - line 5 see page 60, right-hand column, line 7 - line 22</p> <p style="text-align: center;">--- -/-</p>	1-6



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

### \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

9 October 1997

Date of mailing of the international search report

23. 10. 97

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Burgaud, C

# INTERNATIONAL SEARCH REPORT

Int. Patent Application No.  
PCT/GB 97/01834

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>CHOI C S ET AL: "ANALYSIS AND SYNTHESIS OF FACIAL IMAGE SEQUENCES IN MODEL-BASED IMAGE CODING"</p> <p>IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, vol. 4, no. 3, 1 June 1994, pages 257-275, XP000460758 see page 258, right-hand column, line 27 - page 261, left-hand column, line 13</p> <p>---</p>	1-6
X	<p>TERZOPOULOS D ET AL: "ANALYSIS AND SYNTHESIS OF FACIAL IMAGE SEQUENCES USING PHYSICAL AND ANATOMICAL MODELS"</p> <p>IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 15, no. 6, 1 June 1993, pages 569-579, XP000369961 see page 570, right-hand column, line 34 - page 571, left-hand column, line 19 see page 571, right-hand column, line 16 - line 22 see page 573, left-hand column, line 10 - line 22</p> <p>-----</p>	1-6